

VOLUME 66
NUMBER 11

WHOLE No. 343
1952

Psychological Monographs: General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology*
with the *Psychological Monographs*

HERBERT S. CONRAD, *Editor*

Objective Measurement of Reality-Contact Weakness

By

ALBERT KREINHEDER

Los Angeles, California

Based on a Dissertation Presented to the General Faculty of
the Claremont Graduate School in Partial Fulfillment of
the Requirements for the Degree of Doctor of Philosophy

Accepted for publication, January 17, 1952

Price \$1.00

Published by

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1333 SIXTEENTH ST. N.W., WASHINGTON 6, D.C.

CLAREMONT GRADUATE SCHOOL LIBRARY

Psychological Monographs:
General and Applied

COPYRIGHT, 1952, BY THE

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Objective Measurement of
Reality-Contact Weakness

ALBERT KREINBERGER

Ph.D., University of Chicago

Published by the American Psychological Association, Inc., 1200 16th St., N.W., Washington, D.C. 20036

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1200 16th St., N.W., Washington, D.C. 20036

O
I. A
II. S
III. A
IV. A
V. S
A

T
i
malac
feren
indiv
socie
this
tion
ceptu
social
inner
these

An
is th
separ
action
act a
tive
envir
figur
in-an
ble a
phen

T
to th
their
ciatio
dore
Ellis.
for p
60 R
to S
ton
freed
and
Hosp
ness

OBJECTIVE MEASUREMENT OF REALITY-CONTACT WEAKNESS¹

I. <i>Introduction</i>	1
II. <i>Scoring methods</i>	7
III. <i>Results</i>	11
IV. <i>Implications of findings</i>	19
V. <i>Summary</i>	22
<i>References</i>	23

I. INTRODUCTION

THE purpose of this study is to examine experimentally the view that maladjustment involves a degree of difference between what the maladjusted individual sees as real and what his society sees as real. The investigation of this viewpoint has entailed the construction of instruments for measuring perceptual deviations from objective or social reality and an examination of the inner consistency and predictive value of these instruments.

An assumption underlying the research is that perception cannot be entirely separated from memory, thought, and action—that to a large extent individuals act as they perceive (3, p. 367). The adaptive efforts of the individual toward his environment arise out of the total configuration of forces existing in the person-in-an-environment field. The overtly visible aspects of this configuration and the phenomenal aspects (perception) are at

the instant of action in perfect relationship since they are but different ways of viewing the same events. From this point of view, therefore, behavior at odds with the demands of the objective environment coexists with perceptions that ignore, distort, or circumvent the real world as most people understand it.

A. DEFINITION OF TERMS

Defined in terms of the procedures of the study, reality denotes those objectively apparent, verifiable relationships which are generally accepted as self-evident. Strength of the reality contact denotes, in turn, the degree to which an individual's perceptions conform to this normative standard. Strong reality contact is evident in any instance of accurate perception of objective stimuli or in instances of objectively oriented behavior. Weakness of the reality contact may be inferred when individuals are unable to perceive objective stimuli accurately—when, despite deliberate attempts at objectivity, experience and behavior are organized with reference to subjective factors and in disregard of objective fact. Experience, of course, is always subjective, and no weakness of reality contact is implied in the investing of objects with individual significance. Good reality contact merely demands that the raw facts and intractable relations of objective data be not ignored.

¹ The writer wishes to express his appreciation to the members of his doctoral committee for their helpful criticism and for the close association enjoyed with them: Professors F. Theodore Perkins, Florence Mateer, and Robert S. Ellis. He is also indebted to Dr. Bruce T. Jewell for performing the exacting task of ranking the 60 Rorschach records in order of adjustment, to Superintendent Otto L. Gericke of the Patton State Hospital, Patton, California, for the freedom granted in the use of hospital facilities, and to William O. Walcott of the Patton State Hospital Psychology Department for his helpfulness at many times.

It is assumed that the real world may be accurately experienced and that the level of accuracy of one's experience of the real world may be quantitatively described. Bold as such assumptions may appear, they are nothing more than corollaries to the definition of reality which is employed. Whereas absolute reality, if such exists, is nonverifiable, the reality dealt with is a matter of statistical averages and can be known. Objective reality becomes a reference point, and a definable distance exists between this reference point and the point representing personal reality. The shorter this distance, the greater the strength of the reality contact. Reality contact is held to be a quantitative feature of all behavior, cutting across all qualitative dimensions; and, as such, it may furnish a needed conceptual tool for assessing the degree of pathology in various behavioral disorders.

B. THE PROBLEM AND METHOD

The directing hypothesis motivating the present study may be stated as follows:

A determination of the degree of discrepancy between personal reality and objective reality yields a measure that is relevant to an individual's total adjustment.

Specific subquestions derived from this hypothesis have been put to experimental test. These are as follows:

1. Are the reliabilities of the instruments (the several tests of reality contact used in the study) sufficient to warrant further exploration of their value as measures of the strength of the reality contact?
2. Are the validities of the instruments sufficient to warrant further exploration of their value as measures of the reality contact?

3. Are the measures obtained by means of the instruments related to a criterion of the general efficiency of the personality adjustment?

Rorschach results and psychiatric diagnoses were used as criteria of the level of personality adjustment. To insure an adequate test of the questions there were two groups of subjects, the one group composed of paranoid schizophrenics who may be expected to show reality contact weakness in clear form, the other group consisting of normal and psychoneurotic subjects. The latter group, which is called the normal-neurotic group, was made as heterogeneous as possible with respect to level of adjustment. The object of this heterogeneity was the creation of an experimental situation that might reveal a relation between test scores and the level of adjustment within the normal-neurotic range. Analysis of results included intergroup comparison, analysis on the level of both groups combined, and analysis on the level of the normal-neurotic group.

C. SUBJECTS

1. The Normal-Neurotic Group

The subjects comprising the normal-neurotic group were 30 in number and conformed to the following criteria:

1. Young adults of both sexes with ages from 20 to 40.
2. Superior intellectual potential as evidenced by vocational status, educational attainment, or the production of one or more Rorschach responses of high form level.
3. Persons who as a group were as heterogeneous as possible in the efficiency of personality adjustment.
4. The absence of clear-cut psychotic behavior at the time of testing or at any time in the individual's past.

The members of the normal-neurotic group were relatively homogeneous in social and economic background and in educational level. Fifteen of the subjects were college graduates and all but two of the remaining had attended college. Adjustment ratings made from the Rorschach by means of the Munroe revised check list (7) indicated that the attempt to secure heterogeneity on the adjustment dimension was successful.

2. The Psychotic Group

The subjects comprising the psychotic group were 30 in number and conformed to the following criteria:

1. Young adults of both sexes with ages from 20 to 40.
2. Persons who were institutionalized and had been diagnosed as paranoid schizophrenics.
3. Superior intellectual potential as evidenced by educational attainment, by prepsychotic vocational status, or by the production of one or more Rorschach responses of high form level.

In all instances there was a history of bizarre behavior including severe departures from reality demands. The severity of such symptoms at the time of testing varied considerably among the 30 subjects. However, with the exception of one patient, none of the subjects was functioning well enough to be under consideration for discharge from the hospital.

Although all members of the psychotic group had been diagnosed as paranoid schizophrenic, many of the subjects presented mixed symptoms and several—because of personality changes since the time of diagnosis—fit more closely into some other entity. Since paranoid subjects were selected solely for the reason that they tend to be less deteriorated and therefore easier to test, it was not neces-

sary to verify diagnoses. It was essential merely that the presence of psychosis be established, and this, of course, was easy to do.

D. THE INSTRUMENTS AND THEIR MANNER OF ADMINISTRATION

Six tests were administered in three different sessions of approximately one to two hours' duration separated by an interval of at least one day. New tests which were devised for the experiment are reproduced in the original dissertation (5). The sequence of testing was as follows:

- First session: Rorschach
Digit Test, first administration
- Second session: Information Test
Action Picture Test
- Third session: Multiple Choice Stories Test
Picture Description Test
Digit Test, second administration

1. The Rorschach

The Rorschach was used, as described below (II, A), as a basis for securing the criterion of adjustment. It was administered according to the procedures developed by Klopfer and Kelley (4).

2. The Digit Test

The purpose of the Digit Test was to obtain a measure of the perception of the self object by requiring subjects to report upon the accuracy of their own performances. The Digit Test made use of the stimulus material of the digits subtest of the Wechsler-Bellevue Adult Intelligence Scale (9, p. 177). Only one trial was given at each level of difficulty instead of the two trials provided by Wechsler. After repeating each digit series, the subject was asked to report upon the ac-

curacy of his own performance. The preliminary instructions were as follows:

I am going to say some numbers and when I am through I want you to repeat them in the same order as I have said them. After you have repeated the numbers, I shall ask whether you have repeated them correctly or incorrectly. Or perhaps you will be doubtful. You will merely say correct, incorrect, or doubtful.

Before presenting each digit series forward the experimenter said:

Listen carefully and be sure to say the numbers in the same order as I do.

After the subject had repeated each digit series he was asked:

Did you say them correctly, incorrectly, or are you doubtful?

Before presenting the digits backwards, the experimenter explained as follows:

Now I am going to say some numbers and when I am through, I want you to say them backwards. For example, if I say 2, 8, 3, you will say 3, 8, 2. As before, I shall ask whether you said the numbers correctly, incorrectly, or whether you are doubtful.

Before presenting each digit series backwards, the experimenter said:

Listen carefully and be sure to say the numbers backwards.

As before, after each response, the subject was asked to report whether his response was correct, incorrect, or whether he was doubtful.

3. *The Information Test*

The purpose of the Information Test was to provide a situation that disclosed the extent to which subjects were aware of what facts they know and what facts they do not know. In responding to the test items a subject made judgments concerning the conformity of his own beliefs to the standard of objective fact. The test aimed therefore to make manifest, in a limited sense, the manner of perceiving the self object and the congruity of

this view of self with a more objective view.

The Information Test consisted of 50 items of the multiple-choice type with four possible choices in each case. In devising the test an effort was made to secure questions from a wide variety of subject fields and to eliminate in this way the possible influence of specialized knowledge and abilities. Questions were chosen which were judged to be familiar to and yet difficult for the average college graduate.

The preliminary instructions to the subject were as follows:

This is a test to show how much information you know. But it also has another purpose. It shows how aware you are of what you know and what you do not know. Some people believe they know things when they really don't. In taking this test you are supposed to answer only those questions that you are absolutely sure you know. Do not guess. If you are not positive about the answer to a question, leave that question blank. After you are all finished you will have a chance to raise your score by guessing at the questions you left blank. Remember, make as good a score as you can, but don't guess.

The exact wording of the instructions varied somewhat from subject to subject, but the substance was repeated and emphasized until there was no doubt that the subject understood that he was not to guess. After the subject responded to all those items whose answers he was positive he knew, he was furnished with a red pencil and was asked to fill in all the remaining blanks by making the best guesses he could. On the completed test blank all items were marked, either in black or in red.

4. *The Multiple Choice Stories Test*

The Multiple Choice Stories Test consisted of 13 uncompleted stories, each of which was followed by two possible conclusions. The unfinished stories and

their conclusions were mimeographed in such a way that the subject's choice of conclusion could be indicated by a check. One of the two conclusions was more probable of occurrence and more consistent with the facts given. Although the test does not directly measure the perception of reality, its inclusion was based upon the hypothesis that one's method of organizing reality is revealed by the dimension of the story to which he responds.

The experimenter is indebted to Wolfenstein for the idea of the Multiple Choice Stories Test (10). Four of the stories used have the same plots as stories of Wolfenstein although the exact wording is not the same. Nine additional stories were composed by the experimenter.

At the beginning of the test administration the subject was handed the test material and permitted to examine it for a few moments. Then he was instructed as follows:

This test is made up of a number of very short stories. Each story is unfinished, and two possible ways of ending the story are given. You are to read each story and each of the endings which follow it. Then decide which of the two endings you like better. There are no right or wrong answers. Just be guided by your own likes and dislikes.

After a subject had checked all stories, there was an inquiry, the main purpose of which was to determine the basis upon which the choice was made for each story. Responses to the inquiry were recorded verbatim.

5. The Action Picture Test

The materials of the Action Picture Test consisted of 11 eight by five inch cards, each card having upon it a simple line drawing of a human figure in a posture of muscular strain or in the midst of an uncompleted act. Those pictures depicting action in the direction of a

goal were drawn in such a way that successful achievement of the goal was unlikely.

The subject was furnished with ordinary typing paper and a medium sharp lead pencil. He was given the first card and instructed as follows:

This is a test to show how accurately you can draw what you observe. Try to copy each picture exactly as it is as nearly as you are able. But do not use a straight edge nor measure the pictures in any way. There will be 11 pictures in all. You may use as much paper as you wish. There is no time limit.

The factors of interest were two: The total amount of discrepancy between the stimulus pictures and the reproductions, and the degree to which the reproduced drawings reduced the muscular tension of the figures or made the successful completion of the act more certain. These tendencies toward reduction of muscular tension and movement toward the goal were called *closure tendencies*. A hypothesis under consideration was that such instances of closure are indicators of a factor of empathetic response. By empathetic response is meant participation in the emotions or actions of another to the extent of actually feeling the emotions and actually contributing one's own motor effort to the efforts of another. Should closure tendencies be indicative of empathetic response, it is logical to assume that they correspond to dissociated and uncontrolled empathy. This follows from the nature of the instructions which direct the subject to copy the figures exactly as they are.

6. The Picture Description Test

The Picture Description Test was devised in an effort to examine the degree to which the verbal report of a presented visual stimulus diverges from the objective characteristics of the stimulus. Ten pictures obtained from magazines and

mounted on stiff Manila paper 9×12 inches in size were used as stimuli. The pictures were highly structured in the sense of representing familiar objects and people in meaningful relationships. Each picture was placed upon a standard and exposed to the subject's view for a three-second interval; after a ten-second waiting period, the subject described the picture as precisely and completely as possible.

The instructions were as follows:

I am going to show some pictures to you, one at a time. Each picture will be shown for three seconds. After the three seconds are up and the picture is removed, you will have a ten-second period to think about the picture.

Then I will ask you to describe it to me. This will show your ability in observation. The more you remember the better your score will be. But be sure to be completely accurate. Do not interpret the picture or add things that are not actually shown. Just describe what you have seen and see as much as you can. Remember, three seconds, then wait ten seconds, then tell me everything you remember. Just before I show a picture, I shall say, "Ready," so that you know it is coming.

The card to be shown was placed on the standard and covered with another card. The experimenter said, "Ready," and removed the covering card for a period of three seconds as measured by a stop watch. This rough procedure resulted in an actual exposure time a fraction of a second more or less than three seconds, but the amount of error introduced is not large enough to require any more precise method of governing the exposure time.

II. SCORING METHODS

A. THE RORSCHACH

THE QUANTITATIVE measure which served as the criterion of the degree of adjustment was derived from Rorschach results, but the criterion used for analysis on the level of the normal-neurotic group was differently obtained than the criterion used for analysis on the level of both groups combined.

1. The Munroe Check List Rating

The adjustment criterion used for analysis on the level of the normal-neurotic group was the number of tallies assigned to a Rorschach record on the Munroe revised check list. This is a relatively objective rating system which has been standardized on college students similar in many respects to the members of the normal-neurotic group in the present study. By assigning individuals to one of three levels of adjustment on the basis of the number of check list entries, Munroe obtained a contingency coefficient of .62 between the adjustment ratings and carefully obtained external criteria of adjustment (7).

2. The Order of Merit Rating

For analysis on the level of both groups combined an Adjustment Score was obtained by transmuting into units of amount an order of merit ranking of the 60 records on the basis of the efficiency of the subjects' general adjustment. This was used in preference to the Munroe revised check list because Munroe included no psychotics in her standardization study. It was judged to be a feasible procedure because of great range of adjustment represented by the 60 cases. Ranking of the Rorschach records was done by an experienced Rorschach

worker who had no access to any data other than the Rorschach protocols, the Klopfer scoring of each record, and the age and sex of each subject. Good adjustment was defined for him as freedom from psychoneurotic and psychotic characteristics and the ability to use one's capacities in a socially effective way.

A rank-order correlation of .91 was obtained between the ranking made by the Rorschach worker and a similar ranking made by the experimenter. The 60 records were also scored by means of the Basic Rorschach Score (BRS) of Buhler, Buhler, and Lefever (1), and a ρ of .75 was obtained between the BRS and the ratings by the Rorschach worker. The ranking by the Rorschach worker was used in preference to the other two because it was free from the halo effects that might have influenced the experimenter's ranking and it discriminated more effectively than did the BRS between the normal-neurotics and the psychotics.

B. THE DIGIT TEST

As with the other tests of the battery, the Digit Test was scored in such a way that high scores corresponded with weak and low scores with strong reality contact. In scoring the Digit Test score points were added under the following conditions:

Digit series repeated accurately but reported as incorrect.

Digit series repeated inaccurately but reported as correct.

Digit series repeated accurately but reported as doubtful.

Digit series repeated inaccurately but reported as doubtful.

Discrepancies between responses and reports existed in varying degrees so that precise scoring called for a system of

weighting according to the degree of departure from objective fact. Empirical determination of the proper weightings was impossible under the limitations of the study. However, as with the other tests, a heuristic weighting system based upon a priori considerations and the evidence of the data was adopted. Details of the scoring for the several tests are given in the original dissertation (5).

C. THE INFORMATION TEST

In scoring the Information Test, five different tabulations were made: The number of items marked correctly in black, the number of items marked incorrectly in black, the number of red marks, the number marked correctly in red, and the total correct marks in both red and black. The five different classifications were named and symbolized as follows:

- Positive Rights (*PR*)
- Positive Wrongs (*PW*)
- Guesses (*G*)
- Guess Rights (*GR*)
- Knows (*K*)

The first three of these scores are self-explanatory. The fourth score (*GR*) is not the absolute number of items marked correctly in red, but is the figure obtained after the usual correction for guessing is applied. The fifth score, *K*, is the sum of *PR* and *GR*. *K* represents the absolute number of items answered correctly without regard to whether the subject is certain or merely guessing.

The *PW*'s and the *GR*'s were presumed to be instances of reality contact weakness. It is here that the subject was certain of his knowledge when he should have been undecided and undecided when he should have been certain. The sum of *PW* and *GR* could not be used as a measure of reality contact, however, since the absolute size of *PW* and *GR*

varies with the size of *K*. For example, individuals with a very high *K* score will have relatively lower *PW* and *GR* scores; the test is so easy for them that relatively few questions raise any doubt in their minds. Likewise, individuals with very low *K* scores will also tend to have very few *PW*'s and *GR*'s. Many items are completely unfamiliar to them, and such completely unfamiliar items are not ordinarily marked in black nor are they guessed correctly.

How to correct for the size of *K* was the most difficult question in devising the scoring system. The method finally adopted was that which discriminated most significantly between the two groups of subjects. This is expressed by the formula:

$$\frac{PW + GR}{\frac{3}{4}K + g - k}$$

where *g* is the number of guesses (not to exceed 6), and *k* is the amount (not to exceed *g*) by which $\frac{3}{4}K$ exceeds *PW*. In those records where $\frac{3}{4}K$ was greater than *PW* plus *G*, the formula used was:

$$\frac{PW + GR}{PW + G}$$

Scoring weights for several of the test scores, notably the Information Test score (defined above) and the Deviations Score (described in section F, 2, below) were derived from differences noted between the psychotic group and the normal-neurotic group. The resulting weights were then applied to the same groups from which they were derived instead of to a "cross-validation" sample. An inflation of validity coefficients may have been caused by this procedure, but certain facts suggest that such inflation was minimal:

a. Some of the scores (e.g., the Digit Test score and the Multiple Choice Stories Test score) were not weighted by reference to group differences. Yet these scores have validity equal or superior to scores that might be questioned (cf. Tables 3 and 5).

b. Correlations within the normal-neurotic group would not be greatly affected by the

psychotic vs. normal-neurotic group differences that guided the assignment of weights. Yet the r 's within the normal-neurotic group are consistently positive (cf. Table 3).

c. Although the Omissions Score (see section F, 1, below), is less directly affected by group differences than the Deviations Score, it is more highly correlated with the criterion (cf. Tables 3 and 5).

Whatever inflation of validity coefficients may have resulted from the weighting methods used, this inflation is probably not a major factor and does not enter into all scores.

D. THE MULTIPLE CHOICE STORIES TEST

Since the choice of story endings in the Multiple Choice Stories Test was governed by different factors in the case of different subjects, a count of the number of realistic endings chosen has no definite meaning. There is no significant difference between the two groups in the absolute number of realistic endings chosen. The real value of the test lies not in any mechanical summation of the choices but rather in its use with inquiry as a method of eliciting the criteria of choice, the subject's particular mode of organizing or perceiving the problem of the story.

All the bases of choice which were reported in the inquiry were listed, and categories for their classification were determined. The resulting categories were then arranged on an a priori basis into four levels of objectivity. The levels of objectivity and the categories they included are as follows:

Level A: (a) More probable of occurrence; (b) More consistent with the facts given.

Level B: (a) Happier or more satisfactory outcome; (b) Better for characters' personal development; (c) Greater social good; (d) In harmony with moral or ethical principles of a universal kind.

Level C: (a) Rationalizations (subject argues that the pleasant but unrealistic ending is more plausible, logical, or realistic); (b) Personal prejudices, personal likes and dislikes; (c) Consistent with dogmatic, provincial, or childish moral principles.

Level D: The reasons for choice are irrelevant to the central problems of the story. This

includes absurd rationalizations, impossible cause-effect relationships, and the ignoring or twisting of salient relations of the original story.

A ratings were weighted as zero, B ratings as one, C ratings as two, and D ratings as three.

E. THE ACTION PICTURE TEST

Two scores were obtained for the Action Picture Test: viz., a Distortion Score and a Closure Score.

1. The Distortion Score

To obtain the Distortion Score, each drawing was divided into a number of sub-wholes and relationships, and the resulting units were rated piecemeal. Additional points were added when the total gestalt was radically modified.

Each of the component units of a drawing was rated on a five-point scale and assigned values from zero to four. The criteria for assigning the respective values were as follows:

- 0: No clearly visible alteration.
- 1: Alteration is clearly visible but not obvious.
- 2: Alteration is obvious but not marked.
- 3: Marked alteration.
- 4: Extreme alteration.

An alteration, to be scored, had to be a clearly visible alteration in relative size, direction, or proportion. No attention was paid to such factors as the amount of pressure, thickness of lines, erasures, or the like. Examples of subject's drawings and their assigned ratings appear in the original dissertation (5).

After rating the component units of a drawing, the drawing was regarded as a whole to judge whether the essential meaning of the stimulus picture was maintained. In some instances, particularly in the records of psychotics, the gestalt disintegrated so that the reproduction was a meaningless collection of

parts rather than an organized whole. Two degrees of destruction of the gestalt were quite easily recognized. These were called *partial destruction* and *complete destruction*. In partial destruction some semblance of the original unity remains, but only by effort can the relationships be held together. In complete destruction the gestalt is wholly lost. In cases of partial destruction five points were added to the total rating of a picture; and in the case of complete destruction ten points were added.

2. The Closure Score

To ascertain whether closure tendencies were to any extent independent of a general distortion factor, it was necessary to discover whether a subject's tendencies toward closure differed significantly from his tendencies away from closure. The Closure Score was therefore stated in terms of a proportion and was arrived at as follows:

Degrees of error toward and away from closure were rated on a five-point scale according to the same criteria used for rating other alterations in the drawings (II, E, 1). Ratings of errors in the direction of closure were given plus values; ratings of errors in a direction away from closure were given minus values. Instances of partial or complete destruction of the gestalt received a closure rating of minus four. The plus ratings were then divided by the sum of the plus and minus ratings to obtain the Closure Score. In summing the plus and minus ratings, algebraic signs were disregarded. Examples of drawings with their closure ratings appear in the dissertation (5).

F. THE PICTURE DESCRIPTION TEST

Initial scanning of subjects' reports to the Picture Description Test disclosed two main differences between the two groups of subjects. The psychotics not only tended to give distorted reports, but their reports in many instances were

scantier or omitted details that were reported by the majority of the normal-neurotic group. A scoring system was therefore used that recognized both these dimensions. The two aspects of a record were separately scored, as the Omissions Score and the Deviations Score.

1. The Omissions Score

The Omissions Score was based upon the number of frequently reported items omitted from a record. A check list for determining the Omissions Score was prepared, and items on the check list were checked if the corresponding details in the picture were not correctly reported by the subject being rated. Included in the check list were all details that were correctly reported by 50 per cent or more of the normal-neurotic subjects. Items were weighted according to the frequency of their correct report within the normal-neurotic group.

2. The Deviations Score

The Deviations Score was a measure of the degree to which a subject's report deviated from the objective characteristics of the stimulus pictures. To arrive at a method for obtaining the Deviations Score, a classification of the various types of deviations was made and the number of each type of deviation was tabulated for each group. Except for certain liberties taken in the interest of consistency, scoring weights were based upon comparative frequencies from one group to the other. For each class of deviation the weighting approximated the ratio of its frequency in the psychotic group to its frequency in the normal-neurotic group. Examples of kinds of deviations and the weighting assigned to them are given in the dissertation (5).

III. RESULTS

A. RELIABILITY OF TESTS

TABLE 1 present the split-half reliability coefficients and the standard deviations of the several test scores. Except in the case of the Closure Score, the reliabilities are above .60, and the tests may therefore be considered adequate, in so far as reliability is concerned, for the group comparisons made in the present study. Because of the low reliability of the Closure Score, it was excluded from all further analysis.

For individual comparisons, reliability should be .90 or higher; and for practical usefulness the criterion of .90 should be attained within the normal-neurotic range. The increases in test length required by the Spearman-Brown formula to attain the criterion of .90 within the normal-neurotic range are entirely feasible for all tests except the Digit Test and the Information Test. These tests, if increased to the predicted length, would need to be administered in several sessions in order to exclude the factor of fatigue. In addition to increasing test length, reliabilities might be further raised by the correction of certain weaknesses discussed below (III, B, 1). It may be concluded that, with the exception of

the Closure Score, the reliability coefficients are sufficiently high to warrant the further use and development of the tests.

B. VALIDITY OF TESTS

Consideration of the test validities was approached from two standpoints. Efforts were made first to determine the possible influence of irrelevant variables upon the test scores. Secondly, an attempt was made to ascertain whether the behavior sampled by a test is representative of all or part of what is meant by reality contact in real life situations. In this latter sense a test is valid in so far as it has some meaning apart from the actual task it imposes.

1. Freedom from Irrelevant Variables

The problem of eliminating irrelevant variables from the test battery was not equated with a search for primary factors. Even though the tests should be freed from the undue influence of irrelevant factors, the functions measured may still partake of several abilities—and these may turn out to be no different from such classic functions as memory, ability in spatial relations, inductive reasoning, and the like. That this should be true does

TABLE 1
RELIABILITY COEFFICIENTS AND STANDARD DEVIATIONS OF TEST SCORES IN THE
NORMAL-NEUROTIC GROUP AND IN BOTH GROUPS COMBINED*

Test	Normal-Neurotic Group		Both Groups Combined	
	<i>r</i>	<i>SD</i>	<i>r</i>	<i>SD</i>
Digit Test	.64	5.6	.83	10.7
Information	.73	29.8	.71	49.1
Multiple Choice	.84	6.3	.89	10.3
Distortion Score	.88	18.2	.93	42.0
Closure Score	.11	20.8	.26	19.2
Omissions Score	.73	8.4	.91	19.7
Deviations Score	.94	22.2	.92	28.8

* (All coefficients have been corrected by the Spearman-Brown formula)

not necessarily imply that reality contact might better be assessed through an isolation and measurement of its constituents. Although reality-testing tasks undoubtedly involve the known mental abilities to a greater or lesser extent, the issue is not whether an individual has these abilities but how well he can utilize them in the service of relating himself to reality. Reality contact is probably not a discrete function in the sense of being independent of other abilities, but it may prove to be a unique function in that its total variance cannot be accounted for by the variance of its known ingredients. The aim in measuring reality contact is to sink a probe into human behavior at an angle different from that of other tests. Though the tests may cut across the same abilities as sampled by other tests, these abilities are sampled in a unique way.

a. The Digit Test. Examination of the Digit Test results indicated that memory span for digits was not excluded as a variable in those subjects who had unusually high memory span for digits. This was because none of the digit series was difficult enough for them to raise any doubt in their minds about the accuracy of their responses. Although this test defect may have resulted in a substantial error of measurement in several cases, the error was necessarily limited to those few subjects who had unusually high memory span for digits. For adequate measurement of such subjects the addition to the test of several longer digit series is necessary.

An important question was whether or not the test score differed appreciably from the memory span for digits. To test this question, a score for the memory span was computed and was correlated with the Digit Test score. An r of .51 was

obtained. When corrected for attenuation, this was raised to .60. This correlation may be spuriously high because of the test defect mentioned above, but it nevertheless indicates that something in addition to memory span was measured. It is to be expected that reality contact should correlate with memory of all kinds and that this particular test of reality contact should correlate highly with memory span for digits. But the relatively low correlation obtained supports the view that having an ability is not necessarily equivalent to being able to use that ability in relating oneself to reality.

Another problem of interest is whether, when memory span is held constant, any significant relationship remains between the criterion of adjustment and the Digit Test score. Memory span for digits correlated with the Adjustment Score to the extent of .38, while the Digit Test score correlated with the Adjustment Score to the extent of .60. Partialing out memory span resulted in a net correlation between the Adjustment Score and the Digit Test score of .52. It is probable, therefore, that the Digit Test score is not only somewhat independent of memory span but also that it is more relevant to an individual's adjustment than is memory span.

b. The Information Test. The results of an item analysis suggested that eight or ten of the Information Test items introduce a factor of reading comprehension by leading the subject into false premises or by suggesting a fallacious answer. Obviously a subject must understand the question before he is able to make a judgment upon his knowledge concerning the question.

Another problem concerned the extent to which the actual amount of information known varied with the Information

Test score. K , which represents the absolute number of items correctly marked, was correlated with the Information Test score and an r of .61 was obtained. Corrected for attenuation this was raised to .81. The result indicates that something more than a person's store of information was measured. The rather high correlation obtained is probably somewhat spurious because of test defects. Subjects with good analytic ability and a large fund of information are more likely to understand those questions which are misleading. Because of this understanding they will have a greater number of accurate responses (higher K) and a lower Information Test score (indicating higher reality contact). To the extent, therefore, that the Information Test score is contaminated with a factor of reading comprehension the correlation is spuriously high.

The relevance to adjustment of the portion of variance in the Information Test score not accounted for by factual knowledge was also investigated. The r between K and the Adjustment Score was .39 while the r between the Information Test score and the Adjustment Score was .63. Partialing out K , a net correlation remained between the Adjustment Score and the Information Test score of .53. There is reason to believe, therefore, not only that factual knowledge and the awareness of one's factual knowledge are nonequivalent, but also that awareness of one's knowledge has more bearing upon general adjustment than does the mere possession of knowledge.

c. The Multiple Choice Stories Test. The subject's response to the Multiple Choice Stories Test was not restricted to one dimension, since within the limits of the material he was free to choose the dimension to which he responded. For

scoring purposes, however, all responses were ordered to the single dimension of objectivity. It was assumed that, whatever other qualities behavior possesses, it also varies in its degree of objectivity or in the degree of reality contact it manifests. Validation of the scoring procedure requires scaling of responses by a number of judges and other experimentation. However, within the scope of the present study empirical validation of the scale of objectivity appeared in the distribution of responses for the two groups. Of all responses made at the highest level of objectivity, the A level, 75 per cent were made by members of the normal-neurotic group. Proceeding down the levels of objectivity, the proportion of responses from the psychotic group increased progressively.

d. The Distortion Score. Irrelevant variables that might feasibly influence the Distortion Score are motor skill and the care or effort expended in the task.

It is doubtful, however, whether motor skill in this test situation can be thought of as independent of reality contact. Although subjects can point out discrepancies between their own reproductions and the original drawing, it is not necessarily to be inferred that their errors arise from motor disability rather than from reality contact weakness.

Under the lenient scoring system used, a person with normal motor control can reproduce the drawings accurately if he perceives them accurately. The reason individuals make errors which they can later detect as errors is probably because visual perceiving (or the mere contemplation of an object) does not result in the same perceptual emergent as does the visual motor effort of drawing the object. The strong involvement of the motor functions makes for a different sampling of the organism's energy. With some subjects the greater participation of the motor apparatus tends to increase the influence of the inner personal organization causing a greater deviation from the stimulus

object. In other individuals the opposite may be true: Involvement of the motor functions produces a product more in harmony with the stimulus object.

Although capacity for motor control probably has only a negligible effect upon the Distortion Score, a factor of effort is undoubtedly involved. By tentative sketching and frequent erasures certain subjects were able to produce a relatively flawless copy of the original. They improved their visual-motor product by checking and rechecking, perhaps by comparing it with a purely visual perception. With tachistoscopic projection and drawing from memory this factor could be more adequately controlled.

To make the most efficient use of visual motor tests, it is necessary to identify their principal qualitative dimensions. Although distortions may be measured in terms of degree, they also vary in kind. Because of the various qualitative possibilities for response, the total distortion score is a rather crude measure. Nevertheless it is a meaningful measure in that it varies in a general way with adjustment.

e. The Closure Score. The dimensions represented in the Closure Score can only be guessed at since the score itself lacked a sufficient measure of inner consistency. It is probable that the Closure Score is the effect of multiple causes and that these causes operate in varying degrees with different pictures and with different subjects. Negative closure errors, for example, might indicate a bending over backward to check impulsiveness or an actual unresponsiveness to the relationships of the drawing. Another complication is that closure involving the human figures may compete with other closure possibilities of a picture so that what is a good figure from a design sense may not be a good figure from the sense of goal-directedness.

f. The Picture Description Test. From the standpoint of the intrusion of irrelevant factors, there seems to be no serious defect in the Picture Description Test. The Omissions Score, which is a measure of frequently seen items not correctly reported, is qualitatively different from the

other measures used in the study. The other scores are chiefly concerned with the degree to which objective events are subjectively distorted. The Omissions Score is a measure of the degree to which one is aware of those objective relationships which are perceived by most people. Obviously, since the scoring system for computing the Omissions Score is based upon deviations from normative standards, the validity of the measure depends upon a scoring system developed from a much more extensive sampling than that of the present study.

2. Wider Applicability

Are the measurements obtained specific to the conditions of measurement or do they have predictive value for other situations? Findings relevant to this question are reported in Table 2. Additional findings, which concern the relationships between test scores and the criterion of adjustment, are reported later (III, C).

Table 2 reports the test means and standard deviations for each group, together with the critical ratios of the differences between the means on each test. It also reports for each test the percentage of the normal-neurotic scores that overlap the median of the psychotic group and the percentage of psychotics' scores that overlap the median of the normal-neurotic group. In addition to these data for individual tests, the same data are given for a Composite Score. The Composite Score represents the average of the *T* scores of the several tests of reality contact (excluding the Closure Score). The last column of the table reports the correlation of each test with the Composite Score.

The findings reported in Table 2 may be summarized as follows:

1. The hypothesis that there is no

TABLE 2

MEAN AND *SD* OF EACH TEST SCORE AND OF THE COMPOSITE SCORE, CRITICAL RATIOS, PERCENTAGE OVERLAP, AND *r* OF EACH TEST WITH THE COMPOSITE SCORE

Score	N-N. Group		Psychotic		CR	% N-N. o'laps Psy.	% Psy. o'laps N-N.	<i>r</i> with Compos. Score
	Mean	<i>SD</i>	Mean	<i>SD</i>				
Digit	9.87	5.58	21.2	11.46	4.80	10	13	.71
Information	76.8	28.6	135.6	61.02	4.71	7	7	.73
Omissions Score	23.5	8.43	52.0	11.94	10.48	0	0	.82
Deviations Score	26.33	22.2	42.0	32.52	2.16	13	33	.54
Distortion	53.0	18.24	99.34	32.6	6.69	0	13	.79
Closure	48.53	20.82	43.0	17.06	1.14	47	27	
Multiple Choice	12.67	6.30	26.8	8.22	7.42	3	10	.87
Composite	44.87	2.66	55.17	6.16	8.31	0	0	

significant difference between the two groups is rejected beyond the 1 per cent confidence level for all test scores except the Picture Description Deviations Score and the Closure Score. The hypothesis is rejected at about the 5 per cent level for the Deviations Score. The difference between the means of the Closure Score is not significantly greater than chance.

2. There is a tendency for scores of the psychotics to overlap the median of the normal-neurotics to a greater extent than normal-neurotic scores overlap the median of the psychotics.

3. The Composite Score shows less overlapping in either direction than the average overlapping of the scores that compose it.

4. All of the tests comprised in the Composite Score correlate with the Composite Score to an extent significantly greater than that due merely to the part-whole relationship involved.

The first finding—the rejection of the hypothesis of no difference between the means of the two groups—supports the hypothesis of wider applicability. The most plausible inference is that the differences noted are due to a relationship between test scores and degree of psychopathology. Any other variables that might logically be related to the test

scores are substantially equal in the two groups.

The finding that the scores of the psychotics tend to overlap the scores of the normal-neurotics to a greater extent than the reverse is true suggests that a high score on a single test is a stronger indicator of pathology than a low score is of the lack of pathology. The overlapping of the two groups may be partially accounted for by errors of measurement, but several of the overlapping scores deviate too greatly to be explained by error of measurement. This fact is not disconcerting since it is to be expected that some psychotics will display greater reality contact at specific times and with respect to specific situations than will certain neurotic individuals. The relatively large overlap and poor differentiating power of the Deviations Score is cancelled out if the Omissions and Deviations Scores are combined. Certain psychotic subjects were able to attain a low Deviations Score by withdrawing from the situation, but in so doing they accumulated a high Omissions Score.

The significant correlations found between individual tests and the Composite Score also support the hypothesis of wider applicability. It is noted that the tests which correlate most highly with the

Composite Score—the Multiple Choice Stories Test, the Omissions Score, and the Distortion Score—are also highest in the other criteria of validity. Their reliabilities are high and they show the greatest differentiating power in terms of critical ratios and degree of overlap.

C. CORRESPONDENCE OF TEST SCORES AND ADJUSTMENT

1. The Normal-Neurotic Group

The results pertaining to the relationship between test scores and the Munroe check list ratings within the normal-neurotic range are presented in Tables 3 and 4. Table 3 reports the intertest correlations within this range. Table 4 presents the successive multiple \bar{R} 's obtained after the addition of each test, according to the Wherry-Doolittle method of test selection. Since, as is seen from the table, the addition of a fifth test lowered the correlation, there was no point in carrying the computation any further.

Conclusions derived from the results reported in the two tables are as follows:

1. All test scores correlated positively with the criterion, though only the Omissions Score showed a statistically significant relationship. The r between the Omissions Score and the criterion is beyond the 1 per cent level of significance.
2. The intercorrelations of the several

TABLE 4
SUCCESSIVE MULTIPLE \bar{R} 's WITH THE ADJUSTMENT CRITERION RESULTING FROM THE ADDITION OF TESTS BY THE WHERRY-DOOLITTLE TEST SELECTION METHOD (Normal-Neurotic Group)

Tests Added	Resulting \bar{R}
Omissions Score	.57
Information Test	.64
Multiple Choice	.67
Deviations Score	.69
Digit Test	.68

tests were for the most part nonsignificant, indicating a high degree of test specificity within this range of adjustment.

3. The r of .57 between the Omissions Score and the criterion, as well as the multiple \bar{R} of .69, support the hypothesis that reality contact is related to adjustment within the normal-neurotic range. Additional evidences of a relationship are reported below (III, C, 1, a).

a. *Relation of tests to the criterion.* It is premature to draw conclusions from such a small group and without a more carefully obtained criterion. However, comparing the multiple \bar{R} of .69 with Munroe's contingency coefficient of .62 between her check list and carefully obtained validating material (7), it would seem very worth while to develop the tests further. The multiple \bar{R} was obtained with only three tests. Although

TABLE 3
INTERCORRELATIONS OF TEST SCORES AND THE MUNROE ADJUSTMENT RATING IN THE NORMAL-NEUROTIC GROUP

Scores	Digit	Omissions	Deviations	Information	Distortion	Multiple Choice	Adjustment
Digit		.067	-.088	.283	-.033	.588	.338
Omissions Score	.067		.295	-.173	.211	-.141	.571
Deviations Score	-.088	.295		-.041	-.076	-.235	.297
Information	.283	-.173	-.041		.084	.261	.219
Distortion	-.033	.211	-.076	.084		.276	.292
Multiple Choice	.588	-.141	-.235	.261	.276		.228
Adjustment	.338	.571	.297	.219	.298	.228	

four scores were used, the Omissions and Deviations Scores are both obtained from the Picture Description Test. The three tests which produce the multiple \bar{R} are all relatively objective and are adaptable to group administration.

The low correlations between the criterion and all measures except the Omissions Score may mean in part that the functions measured by the tests have very little relationship to adjustment within the normal-neurotic range—although, as already shown, they do have power to differentiate between psychotics and non-psychotics. It is probable that some of the dimensions contained in the tests are related in an inverse or curvilinear way to adjustment within the normal-neurotic range. Curvilinearity could not be established, but analysis of individual records made evident that certain subjects achieved strict adherence to objectivity by special effort. They clung to reality with a tenacity resembling a neurotic defense mechanism. However, when real perceptual disturbances are present, the drive for objectivity does not actually produce objectivity.

Although test scores were frequently relatively lower (stronger reality contact) than the adjustment rating, scores conspicuously higher than the adjustment rating rarely occurred. This was demonstrated statistically in the following way:

The highest test score of each subject was correlated with the adjustment criterion. The resulting r was .58.

The lowest test score of each subject was correlated with the adjustment criterion. The resulting r was .25.

This evidence suggests that low scores are less indicative of good adjustment than high scores are symptomatic of maladjustment. The actual test score that is high does not matter. A real weakness in reality contact at any point is indicative of poor adjustment. It must be borne in mind, however, that the difference between the coefficients of .25 and .58 is significant at only about the 10 per cent level; for verification a larger sample is needed. Nevertheless, the r of .58 is

highly significant and is further evidence of a real relationship between reality contact, as measured, and adjustment within the normal-neurotic range.

There was also further evidence of the relation of test scores to adjustment. When separate tests were combined into a composite score, the correlation with the criterion was increased. Thus, the average of the T scores of five tests (excluding the Omissions Score and the Closure Score) gave the significant r of .53 with the criterion, even though none of the scores contained in this average had a greater r than .34 with the criterion. To predict adjustment within the normal-neurotic range, it appears necessary either to combine a number of specific tests or to construct tests that sample a larger area of the phenomenal field. The success of the Omissions Score is probably partially due to the fact that it presents a more complex reality-testing task.

b. The Omissions Score. Since the Omissions Score is related to the criterion in a significant way, the nature of this measure is important to understand. The Omissions Score represents the number of frequently reported items in the Picture Description Test that are not correctly reported by a subject. It is sensitive both to a narrowing of the perceptual field and to tendencies to organize the perceptual field in unique ways. Since the Omissions Score correlates very highly with the total number of items correctly reported (r equals .87), the main factor measured is probably the degree of constriction of the perceptual field. The theoretical significance of this finding is in its consistency with those theories which suggest that when the organism's capacities are reduced by injury, disease, or senescence, it maintains equilibrium by return to a lesser level of organiza-

TABLE 5
INTERTEST CORRELATIONS IN THE COMBINED GROUP

Scores	Digit	Omissions	Deviations	Information	Distortion	Multiple Choice	Composite	Adjustment
Digit								
Omissions Score	.55			.54	.44	.54	.71	.60
Deviations Score	.08	.26		.53	.63	.59	.82	.66
Information	.54	.53	.13		.36	.47	.54	.45
Distortion	.44	.63	.35	.36		.55	.73	.63
Multiple Choice	.54	.59	.47	.55	.61		.79	.55
Composite	.71	.82	.54	.73	.79	.87		.73
Adjustment	.60	.66	.45	.63	.55	.73	.80	

tion. Because the organism cannot cope with intruding stimuli, it withdraws or incapsulates and only in this way maintains equilibrium. The process has been referred to by Goldstein as the "catastrophic reaction" (2, pp. 117, 136, 318), by Snygg and Combs as "tunnel vision" (8, p. 125), by Lewin as "encysting" (6, p. 94).

2. Both Groups Combined

Table 5 presents the intercorrelations of the several test scores, the Composite Score, and the Adjustment Score in the combined (normal-neurotic and psychotic) group. Comparison of these data with the analogous data of the normal-neurotic group (Table 3) discloses tendencies concerning test specificity and generality:

a. The higher order of interest correlations in the combined group suggests a greater degree of generality for the tests when the analysis covers the total range of adjustment.

b. The higher order of correlations

between tests and criterion in the combined group suggests that adjustment is dependent in a broad way upon reality contact even though minor differences in reality contact may not be directly reflected in adjustment.

c. The greater correspondence of the adjustment criterion with the Composite Score than with any of the individual test scores supports the finding in the normal-neurotic group that the more generalized the measure of reality contact the stronger is its relation to adjustment.

Although the significance of these tendencies is doubtful, particularly since two unlike groups are combined, the conclusion may be tentatively drawn that test scores operate in a fairly specific way within a limited range of adjustment, but that an individual's level of adjustment sets certain broad limits above and below which test scores cannot go. The implications of this conclusion are discussed more fully below (IV, A, 1).

IV. IMPLICATIONS OF FINDINGS

A. THEORETICAL SIGNIFICANCE

1. General Discussion

THE evidence for intra-individual variability in test scores and therefore a certain amount of specificity of reality testing functions is quite clear. Nevertheless this intra-individual variability is relatively small when compared with the range of test scores in the total sample. The degree to which the scores of the psychotics overlapped the scores of the normal-neurotic group and vice versa was relatively small. It may be tentatively concluded, therefore, that some general factor present in the tests of the present battery had a limiting effect upon the range of scores produced by a single individual. The finding of limited intra-individual variability is in line with the idea that specific kinds of reality contact are differentiated parts of the phenomenal field and as such are derived from and dependent upon the total state of the phenomenal field. The argument that these results can be explained just as easily by assuming common elements in the several tests cannot be disposed of, however, without further evidence.

In responding to the tests of the present battery (excepting the Multiple Choice Stories Test) the subject presumably attempts to adapt his own subjective experience as closely as possible to the requirements of objective reality. On highly specific tests a subject's phenomenal experience may be much closer or much farther from objective reality than his average performance, but on tests that are more general in nature, his performance will naturally be nearer to his own average level. To measure a subject's "general" or "average" reality con-

tact weakness, the mean of a number of specific reality contact measures may be taken, or tasks can be contrived which sample a larger area of the phenomenal field. The Picture Description Test, which involves the perception of people and familiar objects in meaningful relationships, probably samples a large area of the phenomenal field. The same is true of the Multiple Choice Stories Test which calls for a consideration of real-life problem situations.

Testing reality contact weakness in specific situations may be conceived of as measuring the distance between the point representing objective reality and the point on the reality dimension at which the task is actually performed. We assume that the subject tries to order his experience as close to the point of objective reality as he can. The validity of this assumption is somewhat questionable, for although the subject is instructed to be as accurate as possible, the task is not crucial to his existence, and he undoubtedly permits a measure of laxness. Some subjects, however, apparently do regard the tasks as crucial to their existence (an unrealistic attitude) and attack them with deadly seriousness. This difference in attitude is usually revealed by internal evidence. It has the greatest effect upon the Distortion Score and the least effect upon the Omissions Score.

With the Multiple Choice Stories Test it cannot be assumed that the subject is even trying to order his experience to the demands of objective reality. In this test the subject is not told to be accurate or realistic but is permitted to seek his own most comfortable point on the reality dimension. The test is therefore useful in determining how close to objective

reality the subject orders his experience when left to his own devices. A strict adherence to objective reality under these conditions is not a favorable sign but appears to be the response of a cramped, prosaic, and overconforming individual. Under these conditions of free choice, however, there is a strong tendency for the level of reality chosen to bear a consistent relation to the person's "general" or "average" reality contact weakness (r between the Multiple Choice Stories Test and the Composite Score equals .87).

2. *Reality Contact in Neurosis*

The frequent assumption that there is no difference in perceptual ability between normals and psychoneurotics is not upheld by the results of the investigation. Four different kinds of evidence point to a relationship between perceptual contact with reality and the degree of adjustment within the normal-neurotic range. These include: (a) the multiple \bar{R} ; (b) the r with the Omissions Score; (c) the r with the highest score of each subject; (d) the r with the average of five scores which by themselves showed no significant relationship with the criterion (III, C, 1).

The general inability in the past to establish the presence of perceptual disturbances among psychoneurotics has probably been due either to the use of methods that are not sensitive enough to reveal the disturbances or to the use of measures that are too specific. Differences between normals and psychoneurotics in the perceptual contact with reality are usually too slight to be evident under ordinary conditions of observation. The stimulus must be reduced to a minimum by tachistoscopic methods or the task must be made difficult in some other way. It is evident also that a large area of the phenomenal field must be sampled. Al-

though individual, specific measures show very little relation to adjustment within the normal-neurotic range, when such specific measures are combined into a composite score a significant relationship appears. The positive results obtained with the Omissions Score are probably due both to the limited exposure method used and to the fact that it samples a larger area of the phenomenal field.

B. APPLICATIONAL SIGNIFICANCE

The idea that reality contact can be regarded as a quantitative feature which cuts across all qualitative dimensions of behavior is upheld to a large extent by the findings of the study. Reality contact undoubtedly fluctuates under differing conditions, but the data of the study nevertheless suggest that a concept of general reality contact, even though it be abstracted from particular instances, can probably furnish a conceptual tool much needed for an appraisal of the extent of maladjustment. Although the investigation of specific kinds of reality contact should not be neglected, the more generalized measures of the present study were shown to have the greatest predictive value. Both in the normal-neurotic group and in the combined group, the composite measures showed the greatest correspondence with the criterion of adjustment. Also, those tests which correlated most highly with the Composite Score (the Multiple Choice Stories Test and the Omissions Score) were most strongly related to the criterion of adjustment.

The adjustment criterion, since it was dependent upon a variety of Rorschach factors, should be expected to correlate most highly with other generalized measures. Nevertheless, obtaining significant

correlations between a general reality-contact score and a general Rorschach adjustment score implies a validity for both measures. If general adjustment were not a valid concept to begin with, the validity of a general reality-contact score could not be demonstrated. And although specific areas of maladjustment must be located in order to apply psychotherapeutic correctives, the practical value of a general estimate of adjustment efficiency is not to be denied.

The battery in its present form is adequate for distinguishing psychotics from neurotics in all but a few borderline cases, but it is not sufficiently sensitive to neurotic disturbances to make valid conclusions regarding a particular individual. A rather maladjusted person will occasionally produce a low Composite Score (indicating strong reality contact) so that, if the battery were to be used for screening purposes, certain malfunctioning individuals would go undetected. However, in the present data, a conspicuously high Composite Score always indicates severe disturbance of some sort.

Reality-contact measurement may supplement but is unlikely to supplant the usual diagnostic techniques. Practical advantages include the objective nature of the tests and their adaptability to group

use. Should the possibilities suggested by the present data be fulfilled in later experimentation, a way to gauge the extent of maladjustment will be established. Yet—whether or not the relationship between maladjustment and reality contact is universally valid—psychodiagnosis obviously implies more than a determination of the degree of malfunction. Also, it is wholly probable that in certain maladjustments—perhaps in the affective disorders—real pathology may exist without significant reality-contact loss. On the other hand, perceptual deviations may in such cases be merely more specific or less observable.

Although not reported in this paper, the data of the present investigation also suggest that specific kinds of reality-contact weakness are associated with specific behavioral tendencies. Several hypotheses along this line are offered in the original dissertation (5). Knowledge of whether reality-contact measures can function in precise personality description and in differential diagnosis depends upon the testing of such hypotheses. And this endeavor depends in turn upon well-developed, specific tests and criteria, which should, perhaps, be closer to actual-life reality than are the laboratory-type exercises of the present battery.

V. SUMMARY

FIVE tests were devised for the purpose of measuring weakness of the reality contact and were administered to a group of 30 paranoid schizophrenics and to a heterogeneous group of 30 subjects, both normal and psychoneurotic. Seven scores were derived from these five tests. The Rorschach was also administered and was used as a basis for estimating the general level of adjustment of each subject. Questions to be investigated were the reliability and validity of the several tests of reality contact, and the relation of the measures of reality contact to the level of adjustment.

A. RESULTS

1. The split-half reliability after correction by the Spearman-Brown formula was sufficiently high to justify use of all scores, except the Closure Score of the Action Picture Test, for group comparison. The Closure Score was therefore omitted from all further analysis. The reliability coefficients of the remaining tests varied from .64 to .94 in the normal-neurotic group, and from .71 to .93 in both groups combined (normal-neurotic plus psychotic).

2. With respect to the freedom of the tests from irrelevant factors, data were insufficient to give a clear-cut answer. It was shown, however, that the Digit Test measured something in addition to memory span for digits, and that the Information Test measured something in addition to the actual amount of information known. An irrelevant variable that appeared to influence the score on the Information Test was the factor of reading comprehension.

3. Significant differences between the mean of the normal-neurotic group and

the mean of the psychotic group were obtained in all scores except the Closure Score. The critical ratios of these significant differences varied from 2.16 to 10.48. Several of these figures may have been inflated, though probably not appreciably, by the method of weighting scores (II, C).

4. Significant relationships were found between the measures of reality contact and the degree of adjustment within the normal-neurotic range. The r between the Omissions Score and the criterion of adjustment equaled .57. A multiple R representing the correlation between the criterion and three tests (four scores) equaled .69. The r between the highest score of each subject and the criterion was .58. The r between the average of five scores (excluding the Omissions Score) and the criterion was .53, even though none of the scores contained in this average correlated with the criterion to an extent greater than .34.

5. Although there was a large degree of generality or correlation among the several tests in the full-range sample (both groups combined), in the normal-neurotic group the tests were much more specific (uncorrelated) in their operation.

B. CONCLUSIONS

1. The experimental results support the belief that relatively objective and reliable tests of reality contact can be devised.

2. The results of the study support the view that such tests possess a validity that extends beyond the specific conditions of measurement.

3. The data of the study uphold the idea that the weakness of the reality contact is related to the general level of maladjustment, both within the normal-

neurotic range and in the total range of adjustment.

4. Within the limits of the study, reality contact appeared to be both specific and general in its operation. Although a certain amount of intra-individual variability existed from one reality testing task to another, the intra-individual variability was relatively small when compared to the range of test scores within the total sample.

5. The evidence of the study supports the idea that a concept of general reality contact has practical usefulness even though it be abstracted from specific instances.

6. The generally positive results of the experiment favor the further refinement of the instruments used, the development of new instruments, and the further investigation of the dimensions and correlates of reality contact.

REFERENCES

1. BUHLER, C., BUHLER, K., & LEFEVER, D. W. *Development of the basic Rorschach score with manual of directions*. Los Angeles: Mimeographed by authors, 1949.
2. GOLDSTEIN, K. *The organism. A holistic approach to biology derived from pathological data in man*. New York: American Book Co., 1939.
3. HELSON, H. The psychology of gestalt. *Amer. J. Psychol.*, 1925, 36, 342-370.
4. KLOPPER, B., & KELLEY, D. M. *The Rorschach technique*. New York: World Book Co., 1942.
5. KREINHEDER, A. *The strength of the reality contact, an attempt at its objective measurement*. Unpublished doctoral dissertation, Claremont Graduate School, 1951.
6. LEWIN, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1935.
7. MUNROE, R. L. *Prediction of the adjustment and academic performance of college students by a modification of the Rorschach method*. Stanford, Calif.: Stanford Univer. Press, 1945.
8. SNYGG, D., & COMBS, A. W. *Individual behavior*. New York: Harper, 1949.
9. WECHSLER, D. *The measurement of adult intelligence*. Baltimore: Williams & Wilkins, 1944.
10. WOLFENSTEIN, M. The reality principles in story preferences of neurotics and psychotics. *Charact. & Pers.*, 1944, 13, 135-151.